



# Demographic history and spatial genetic structure in a remnant population of the subtropical tree *Anadenanthera colubrina* var. *cebil* (Griseb.) Altschul (Fabaceae)

Alejandra L. Goncalves<sup>1,2,3</sup> · María V. García<sup>1,2,3</sup> · Myriam Heuertz<sup>4</sup> · Santiago C. González-Martínez<sup>4</sup> 

Received: 19 July 2018 / Accepted: 2 January 2019 / Published online: 13 February 2019  
© INRA and Springer-Verlag France SAS, part of Springer Nature 2019

## Abstract

• **Key message** A remnant population of *Anadenanthera colubrina* var. *cebil* in Northern Argentina showed a mixed mating system, high genetic diversity, and moderate spatial genetic structure, which was stronger in saplings than in adults. Demographic history analyses revealed an ancient population expansion. Despite high genetic diversity, high inbreeding suggests caution in the use of this stand as seed source.

• **Context** Information on fine-scale spatial genetic structure (FSGS) and demographic history is essential to determine which mechanisms are responsible for population persistence and evolution. This is particularly important in fragmented biomes, such as the seasonally dry tropical forests.

• **Aims** To assess the level of genetic diversity and population genetic structure in a remnant population of *A. colubrina* var. *cebil*, and to evaluate the influence of historical and contemporary environmental change on the genetic constitution of this population.

• **Methods** Eight microsatellites were typed in 60 adults and 59 saplings. The existence of (non-spatial) genetic clusters was evaluated using STRUCTURE and PCAs. FSGS was evaluated by kinship analyses and sPCA. MCMCglmm models were used to provide insights into factors underlying FSGS. Demographic history was studied using bottleneck statistics and approximate Bayesian computation.

• **Results** We found high levels of genetic diversity and high inbreeding. Genetic structure was stronger in saplings than in adult trees, probably due to assortative mating, and was not explained by altitude or DBH. Demographic analyses suggested an ancient population expansion.

• **Conclusion** Patterns of inbreeding and relatedness suggest a mixed mating system. High genetic diversity and moderate genetic structure suggest long-term population viability. High inbreeding suggests caution when using this stand as a source of material for reforestation.

**Handling Editor:** Bruno Fady

## Contribution of the co-authors

Conceptualization: MVG; Methodology: MVG, MH and SCGM; Formal Analysis: ALG, MH; Writing – original draft: ALG; Writing – review & editing: ALG, MVG, MH and SCGM; Supervision: MVG and SCGM; Project Administration: MVG; Funding acquisition: MVG.

✉ Santiago C. González-Martínez  
santiago.gonzalez-martinez@inra.fr

<sup>1</sup> Facultad de Ciencias Exactas, Químicas y Naturales, Universidad Nacional de Misiones, Posadas, Misiones, Argentina

<sup>2</sup> Instituto de Biología Subtropical – Nodo Posadas (UNaM – CONICET), Posadas, Argentina

<sup>3</sup> Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina

<sup>4</sup> BIOGECO, INRA, Univ. Bordeaux, 33610 Cestas, France

**Keywords** Paranaense biogeographic province · Fine-scale spatial genetic structure · Demographic modelling · Genetic diversity · Inbreeding · Microsatellites

## 1 Introduction

The knowledge of spatial patterns of genetic diversity is fundamental for the conservation of native forest tree species since it can inform the design of conservation strategies (Frankham 2003; Escudero et al. 2003; Davies et al. 2013). Fine-scale spatial genetic structure (FSGS), defined broadly as the non-random spatial distribution of genotypes, has been described in many plant populations (see reviews in Vekemans and Hardy 2004; Troupin et al. 2006; Hardy et al. 2006; Dick et al. 2008; Buzatti et al. 2012; Berens et al. 2014; Ewédjè et al. 2017). FSGS is shaped by the complex interplay

of microevolutionary processes (i.e., genetic drift, gene flow, and natural selection) and mating system (Epperson 2003; Dick et al. 2008). In non-equilibrium populations, demographic history can also affect FSGS patterns: for example, locally co-occurring gene pools or allele frequency gradients can result from multiple waves of colonization and population establishment or from habitat-mediated selection (Hampe et al. 2010; Torroba-Balmori et al. 2017). Therefore, information on the demographic history is also essential to determine which mechanisms are responsible for generating the temporal dynamics of FSGS (Jones and Hubbell 2006).

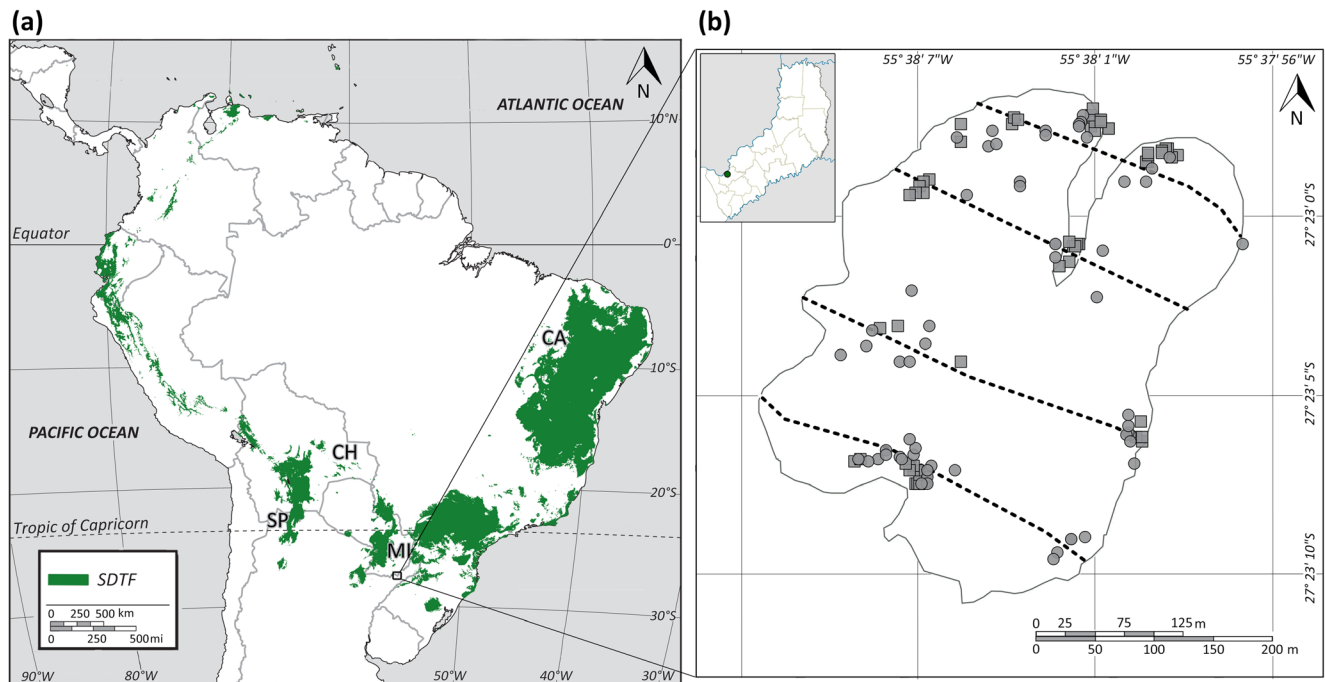
The fragmentation of formerly continuous tree populations disrupts natural ecological and evolutionary processes, and can adversely modify their genetic composition (Hamrick 2004). Decrease in forest fragment size, together with increased fragment isolation, habitat loss, and reduction in tree densities, may lead to genetic erosion due to the effects of genetic drift and inbreeding, and lower gene flow among forest remnants (Young et al. 1996; Aguilar et al. 2008; Fageria and Rajora 2013). These effects are exacerbated when reductions in population size occur in populations that display spatial genetic autocorrelation (Nason and Hamrick 1997; Aldrich et al. 1998).

In long-lived tree species, changes in effective population size, defined as the average number of individuals that contribute offspring to the next generation (Ridley 2003), may be slow and difficult to detect. Moreover, the fact that genetic processes operate in time units of generations may introduce important time lags in the response of forest trees to landscape change, including fragmentation (Wagner and Fortin 2013). Time lags between landscape change and subsequent genetic change can vary from a few to several thousand generations depending on the type of landscape change and the specific genetic parameters studied (Epps and Keyghobadi 2015). For example, after a change in subpopulation connectivity, measures of inbreeding or heterozygosity move towards equilibrium much more slowly than measures of genetic differentiation (Epps and Keyghobadi 2015, and references therein). Nevertheless, the comparison of the spatial pattern and genetic structure across different life stages has proven to be a useful tool to study the ecological and evolutionary processes underlying population dynamics in forest trees (see examples in Jones and Hubbell 2006; Berens et al. 2014). In forest trees, saplings are usually spatially clustered and adult trees are more randomly distributed, suggesting assortative mating and/or density-dependent mortality (e.g., Janzen-Connell effects; Clark and Clark 1984; Comita et al. 2014).

Here, we studied the FSGS and the demographic history of a remnant population of the insect-pollinated and gravity-dispersed subtropical tree *Anadenanthera colubrina* var. *cebil* (Griseb.) Altschul, Fabaceae (Spanish: cebil or curupay; Portuguese: angico; also known as “vilca”, which is derived from the Quechua voice “wilka”, meaning “sacred”), in the

Campo San Juan Nature Reserve (Misiones, northeastern Argentina). *Anadenanthera colubrina* var. *cebil* is a long-lived and semi-deciduous canopy tree that can reach up to 35 m in height. It has compound bipinnate leaves with specialized ant glands (extrafloral nectaries), hermaphroditic flowers in inflorescences, and long legume fruits (von Reis Altschul 1964; Justiniano and Fredericksen 1998; Cialdella 2000; Klitgård and Lewis 2010). The mating system of *A. colubrina* var. *cebil* is presumed to be mostly outcrossing (Cialdella 2000), although one study in Ribeirão Preto (São Paulo, Brazil) showed a mixed mating system (multi-locus outcrossing rate,  $t_m = 0.619$ , Feres 2013). Bees are the main pollinators, and seeds are dispersed by autochory or anemochory after pod dehiscence (Justiniano and Fredericksen 1998; de Noir et al. 2002). *Anadenanthera colubrina* var. *cebil* occupies a wide distribution range (Fig. 6 in Annex 4) in the seasonally dry tropical forests (SDTFs), a relatively poorly known biome with a strongly fragmented distribution across Latin America and the Caribbean (Särkinen et al. 2011; Moggi et al. 2015; DRYFLOR 2016). *Anadenanthera colubrina* var. *cebil* has been proposed by Prado and Gibbs (1993) as the most important tree species in this biome, as it is present in 8 out of the 12 floristic groups described for Neotropical SDTFs (DRYFLOR 2016) and occurs in four major SDTF nuclei—Caatinga, Misiones, Chiquitania, and Subandean Piedmont (Särkinen et al. 2011; Fig. 1a). The current distribution of SDTFs in South America is highly fragmented due to both natural factors (e.g., climatic fluctuations during the Pleistocene) and human disturbance (e.g., land use changes and deforestation) (Werneck et al. 2011; Sánchez-Azofeifa and Portillo-Quintero 2011).

In a previous study, Barrandeguy et al. (2014) detected genetic signatures of ancient fragmentation in both nuclear and chloroplast genomes for *A. colubrina* var. *cebil* and suggested (but did not test) that demographic instability may underlie current genetic structure at the large geographical scale. By studying local genetic structure and demographic history in one typical location of the species, we aimed at providing insights into the causes underlying these patterns, which is especially relevant for highly fragmented and poorly known biomes, such as the SDTF. The specific aims of our study are (i) to characterize patterns of genetic diversity and FSGS of *A. colubrina* var. *cebil* in a typical remnant population from the Misiones nucleus in the Paranaense biogeographic province, Argentina; (ii) to investigate one-generation changes in these patterns by comparing adults and saplings from the same population; and (iii) to provide insights into demographic history and the causes underlying population structure in this population. Current levels of protection of SDTFs are inadequate, as protected area is very small with respect to the high diversity, high endemism, and high floristic turnover found in the biome (DRYFLOR 2016). Knowledge on genetic diversity and population genetic structure, in relationship to



**Fig. 1** **a** Current distribution of seasonally dry tropical forests (SDTFs) in South America according to Särkinen et al. (2011). Four nuclei are identified: Caatinga (CA), Misiones (MI), Chiquitania (CH), and Subandean Piedmont (SP). The location of the studied area, Campo San Juan Nature

Reserve, is indicated. **b** Studied forest stand indicating adult trees (circles) and saplings (squares) of *Anadenanthera colubrina* var. *cebil*, and the four transects sampled (dashed lines)

demographical processes in keystone SDTF species, such as *A. colubrina* var. *cebil*, can provide valuable information for in situ dynamic conservation and the establishment of sampling strategies for ex situ conservation and reforestation.

## 2 Material and methods

### 2.1 Study area

Campo San Juan (CSJ) is a Nature Reserve located in the Paranaense biogeographic province (Cabrera 1971) within the Misiones nucleus of the SDTFs (Fig. 1b). Native grasslands and forests cover 90% of the reserve, in approximately equal proportions (Falguera et al. 2015). Within CSJ Nature Reserve, we studied a forest stand of c. 20 ha located in the center of a fragmented forest extending into the grassland matrix (Fig. 1b). Besides *Anadenanthera colubrina* var. *cebil*, the main forest trees in the CSJ Nature Reserve are *Myracrodruon balansae* (Anacardiaceae), *Handroanthus heptaphyllus* (Bignoniaceae), *Cedrela fissilis* (Meliaceae), *Parapiptadenia rigida*, and *Peltophorum dubium* (Leguminosae) (Falguera et al. 2015). Four parallel transects were established (T4, T5, T6, and T7), crossing the stand from the southwestern to the northeastern forest edge (Fig. 1b). Transects were around 340 m long by 50 m wide, covering approximately 40% of the total forest area.

### 2.2 Plant material

Two life stages of *A. colubrina* var. *cebil* were sampled along the four transects. In total, young leaves from 119 individuals were collected, including all reproductively mature trees in the transects (60 individuals), and ~60% of the saplings (59 individuals). Trees were georeferenced by GPS (Geographic Position System) by means of a Garmin eTrex® 20× receiver (precision of  $\pm 3$  m) and identified by an individual code. Diameter at 1.30 m (DBH) was also measured in adult trees. Leaves were dried in salt (Carrió and Rosselló 2013) and stored at room temperature until DNA extraction.

### 2.3 DNA extraction and microsatellite genotyping

Total genomic DNA was extracted from dry leaves using the ROSE extraction protocol (Steiner et al. 1995) modified by García et al. (2007). Eight specific nuclear microsatellite markers (SSR) for *A. colubrina* var. *cebil* were used to genotype all individuals (Barrandeguy et al. 2012). The amplifications by polymerase chain reaction (PCR) were performed in a final volume of 14  $\mu$ l containing 0.5 ng/ $\mu$ l of genomic DNA, 1× Buffer, 2.5 mM  $\text{Cl}_2\text{Mg}$ , 0.175 mM of each dNTP, 0.75 U Hot Start DNA Polymerase, 0.33 pmol of reverse primer, 0.20 pmol of forward primer with M13 tail at the 5'-end, and 0.33 pmol of universal FAM- or HEX-labeled M13 primer, following the method proposed by Schuelke (2000). PCR was performed in a gradient cycler (Biometra) using a touchdown

program or a fixed annealing temperature, depending on the locus, following Barrandeguy et al. (2012). The amplification cycles were divided in two phases: the first consisted of 30 cycles using the annealing temperature of the specific SSR primers and the second of 8 cycles using the M13 primer annealing temperature (53 °C). The capillary electrophoresis was carried out in an ABI 3730XL Genetic Analyzer (Applied Biosystems®) and fragment sizes were scored with Peak Scanner™ v1.0 (Applied Biosystems®) software using 400HD-ROX™ (Applied Biosystems®) as internal size standard.

## 2.4 Data analysis

### 2.4.1 Genetic diversity and inbreeding

Genetic variation in the population was characterized by number of alleles ( $N_A$ ), effective number of alleles ( $N_E$ ), number of private alleles ( $N_{PA}$ ), observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity, and allelic richness ( $R$ ). Observed and expected heterozygosity and allelic richness were computed in SPAGeDi 1.5a (Hardy and Vekemans 2002) whereas GenAlEx 6.5 (Peakall and Smouse 2012) was used for all other genetic diversity estimates. Significant differences among life stages in  $R$ ,  $H_O$ , and  $H_E$  were tested using a permutation procedure (10,000 iterations) in FSTAT 2.9.3.2 (Goudet 1995). A global population test of heterozygote deficit across loci was computed in GENEPOP 4.2.2 (Rousset 2008) to detect departures from Hardy–Weinberg equilibrium (HWE) with the following parameters: dememorization = 10,000; batches = 100; and iterations = 5000. The inbreeding coefficient ( $F_{IS}$ ) was calculated using SPAGeDi 1.5a (Hardy and Vekemans 2002). The presence of null alleles and genotyping errors were estimated in each locus using MicroChecker (van Oosterhout et al. 2004). The individual inbreeding model (IIM), implemented as a Bayesian approach, was used to partition out the influence of null alleles on  $F_{IS}$  values using INest (Chybicki and Burczyk 2009).

### 2.4.2 Genetic clusters and Wahlund effects

The existence of genetic clusters in the population was first explored with a Bayesian model-based cluster analysis using STRUCTURE 2.3.4 (Pritchard et al. 2000). This method estimates the number of genetic clusters ( $K$ ) in the data and estimates the ancestry of individuals in these clusters (Pritchard et al. 2000). Bayesian analysis was performed using the admixture model with correlated allele frequencies and  $K$  from 1 to 8. The model was run separately for adults and saplings with 10 independent simulations for each  $K$ , and a burn-in length of 500,000 and a run length of 750,000 MCMC iterations. The optimal number of clusters was determined using the ad hoc  $\Delta K$  statistic (Evanno et al. 2005), based on

the second order of change in the log likelihood of data ( $\Delta K$ ) as a function of  $K$ , using Structure Harvester 0.6.92 web application (Earl and vonHoldt 2012). Second, principal component analysis (PCA) as implemented in the ADEGENET 2.1.1 package in R (Jombart 2008) was used to validate STRUCTURE genetic clusters. Geographic maps showing the distribution of genetic clusters were drawn using QGIS v2.18.13 (Quantum GIS Development Team 2017).

Based on the optimal number of genetic clusters, individuals were assigned to the cluster in which they had the highest proportion of estimated ancestry, and genetic differentiation (i.e., the distribution of genetic variation within and among clusters) was examined using analysis of molecular variance (AMOVA) in Arlequin 3.5 (Excoffier and Lischer 2010). Statistical significance of AMOVA was calculated using 10,000 permutations. In addition, we estimated unbiased  $F_{ST}$  values using the excluding null allele (ENA) correction method in FreeNA software (Chapuis and Estoup 2007). Genetic differentiation was further tested using Jost's  $D$  index (Jost 2008), which partitions heterozygosity within and among clusters (Jost et al. 2018). Statistical significance was tested by means of 1000 bootstraps using DEMETICS v0.8.7 package in R (Gerlach et al. 2010). In addition, to test for Wahlund effects, inbreeding coefficients ( $F_{IS}$ ), both corrected and uncorrected for the presence of null alleles, were computed for each genetic cluster using the methods described above. A Wahlund effect is observed when inbreeding at the population level is due to population substructure, thus, under a Wahlund effect,  $F_{IS}$  within genetic clusters should be lower than  $F_{IS}$  of the entire forest stand.

### 2.4.3 Fine-scale spatial genetic structure

Spatial genetic structure was first investigated using a spatial principal component analysis, sPCA, in ADEGENET 2.1.1 package in R (Jombart 2008). Strength of spatial genetic structure was estimated by the eigenvalue of the first sPCA axis, eig.sPCA. We tested for global structure, such as allele frequency gradients or spatial autocorrelation, using the  $G$  test, and for local structure, such as a spatial repulsion signal between genetic clusters, using the  $L$  test. Second, spatial autocorrelation analyses based on pairwise kinship coefficients ( $F_{ij}$ ) between individuals (Loiselle et al. 1995) against physical distance were done using SPAGeDi 1.5a (Hardy and Vekemans 2002). Mean multi-locus kinship coefficients ( $F_{(d)}$ ) were computed separately for each life stage by means of pairwise comparisons for eight distance classes with upper bounds of 10, 20, 50, 100, 150, 200, 300, and 500 m, following recommendations by Doligez et al. (1998). Standard errors for the kinship coefficients were estimated using a jackknife procedure over loci. To test for significance of fine-scale spatial genetic structure (FSGS),  $F_{ij}$  for all pairs of individuals was plotted against the logarithm (for two-dimensional space)



of pairwise spatial distance, and significance of the regression slope  $b$  was assessed by 10,000 permutations of the columns and rows of the distance matrix. To compare the extent of FSGS with that in other plant species, we also computed the  $Sp$  statistic, which is defined as  $Sp = -b/(1 - F_1)$ , where  $F_1$  is the mean  $F_{ij}$  between neighboring individuals, i.e., for the first distance class ( $< 10$  m), and  $b$  is the regression slope based on the logarithm of Euclidean geographic distance (Vekemans and Hardy 2004).

Finally, to further investigate the factors influencing FSGS, we used a Bayesian generalized linear mixed modelling approach implemented in the MCMCglmm package (Hadfield 2010) in R. We contrasted seven models using three predictor variables and combinations thereof (as fixed effects), to predict the inter-individual kinship matrix of adult trees only: 1. NULL, a null model without any predictor, 2. GEO, with a predictor matrix representing the Euclidean geographic distance between individuals, computed from latitude and longitude, 3. ALT, with a predictor matrix based on Euclidean distance in altitude only, 4. DBH, with a predictor matrix based on diameter differences at 1.30 m (DBH), assuming that DBH could be considered as proxy for distinct groups of individuals sharing microhabitat or growth strategy, 5. GEO\_ALT, with both geographic and altitude predictor matrices, 6. GEO\_DBH, with geographic and DBH predictor matrices, and 7. GEO\_DBH\_ALT, with the three predictor matrices. The number of MCMC iterations was set to 200,000, after a burn-in of 50,000 and using a thinning interval of 5000. The deviance information criterion (Spiegelhalter et al. 2002), DIC, of each model was recorded and used for model comparison. The model with the lowest DIC was considered to best describe the data.

#### 2.4.4 Demographic history

Demographic history of the forest stand (based on all adult trees) was first explored by computing the “bottleneck” statistic,  $T_2$ , in Bottleneck 1.2.02 (Cornuet and Luikart 1996; Piry et al. 1999). This statistic was estimated using a two-phase mutation (TPM) model, which is the most appropriate for SSRs (Di Rienzo et al. 1994; Piry et al. 1999), by means of 10,000 iterations, and considering 95% single-step mutations, 5% multiple-step mutations, and a variance of 12 among multiple steps. This method is suitable to detect recent bottlenecks of low magnitude (Williamson-Natesan 2005). Positive  $T_2$  values indicate a heterozygosity excess that can be associated with demographic bottlenecks (Piry et al. 1999). The method can also be used to detect population expansions, which are characterized by a heterozygosity deficiency (Cornuet and Luikart 1996). The significance of  $T_2$  was tested using a Wilcoxon sign-rank test. Additionally, the Garza–Williamson  $M$  ratio test was computed in

Arlequin 3.5 (Excoffier and Lischer 2010), where  $M$  is the mean ratio of the number of alleles to the allele size range (Garza and Williamson 2001). A value of  $M < 0.68$  is associated with ancient and severe bottleneck signals (Garza and Williamson 2001; Williamson-Natesan 2005).

Approximate Bayesian computation (ABC) methods were used for more accurate demographic inference (Csilléry et al. 2010), using DIYABC v1.0.4.46beta (Cornuet et al. 2008). Based on historical records (Bauni and Homberg 2015; Werneck et al. 2011), two demographical scenarios were defined for the studied area, modelled through simple two-epoch models of population size change, without fixing the direction of change (i.e., each scenario allowed for both population expansion and population decline). The first scenario assumed a recent demographic change (one to five generations ago) corresponding to strong habitat transformation due to known human activity in the region (Bauni et al. 2013). The second scenario reflected an ancient demographic change (100 to 700 generations ago) associated with the expansion-contraction dynamics described in SDTFs (Prado and Gibbs 1993; Pennington et al. 2000; Werneck et al. 2011). These simple two-epoch models of population expansion/decline were compared with a null model of constant population size, using logistic regression to estimate the posterior probability for each scenario, and to estimate the following parameters: current ( $N_1$ ) and ancient population size ( $N_0$ ), and time of demographic change ( $t$ ). Prior distributions, model parameters, and summary statistics are reported in Annex 1 (Table 4).

### 3 Results

#### 3.1 Genetic diversity and inbreeding

All eight nuclear microsatellite loci were highly polymorphic in both adults and saplings (Table 1). Allelic richness ( $R$ ), and both observed ( $H_O$ ) and expected ( $H_E$ ) heterozygosity did not show significant differences across life stages ( $p$  values of 0.12, 0.57, and 0.27 for  $R$ ,  $H_O$ , and  $H_E$ , respectively).  $H_O$  was significantly lower than  $H_E$  resulting in a high inbreeding coefficient for both adults and saplings ( $F_{IS} = 0.41$  and 0.40, respectively). High  $F_{IS}$  could be explained by the presence of null alleles (Chakraborty et al. 1994), or alternatively, by population substructure (Wahlund effect, see below), assortative mating, and/or selfing (Morand et al. 2002). Despite the high frequency of null alleles ( $r = 0.19$ ), as estimated by the Brookfield equation (Brookfield 1996), corrected  $F_{IS}$  values were similar to those estimated without correction: 0.39 (95% CIs of 0.35–0.44) in adults and 0.35 (95% CIs of 0.28–0.42) in saplings (Table 1).

**Table 1** Genetic diversity and inbreeding in the studied population remnant of *Anadenanthera colubrina* var. *cebil*

Locus	Adults ( <i>N</i> = 60)										Saplings ( <i>N</i> = 59)									
	<i>N</i> <sub>A</sub>	<i>N</i> <sub>E</sub>	<i>N</i> <sub>PA</sub>	<i>R</i>	<i>H</i> <sub>O</sub>	<i>H</i> <sub>E</sub>	<i>F</i> <sub>IS</sub>	<i>r</i>	<i>F</i> <sub>IS</sub> (IIM)	<i>N</i> <sub>A</sub>	<i>N</i> <sub>E</sub>	<i>N</i> <sub>PA</sub>	<i>R</i>	<i>H</i> <sub>O</sub>	<i>H</i> <sub>E</sub>	<i>F</i> <sub>IS</sub>	<i>r</i>	<i>F</i> <sub>IS</sub> (IIM)		
<i>Ac</i> 34.3	28	8.82	12	27.83	0.550	0.887	0.387	0.178		23	7.63	7	22.69	0.534	0.869	0.392	0.182			
<i>Ac</i> 48.1	33	17.48	6	32.83	0.517	0.943	0.459	0.219		38	17.00	12	38.00	0.536	0.941	0.438	0.218			
<i>Ac</i> 11.2	14	4.81	4	13.95	0.450	0.792	0.439	0.191		16	6.13	5	15.79	0.466	0.837	0.451	0.201			
<i>Ac</i> 28.3	36	20.69	9	35.85	0.550	0.952	0.429	0.206		34	18.85	8	33.51	0.672	0.947	0.298	0.177			
<i>Ac</i> 157.1	42	16.44	14	41.70	0.800	0.939	0.156	0.072		39	17.19	11	38.17	0.661	0.942	0.306	0.110			
<i>Ac</i> 41.1	23	10.21	7	22.90	0.333	0.902	0.636	0.299		23	10.28	7	22.58	0.424	0.903	0.537	0.278			
<i>Ac</i> 172.1	26	8.85	6	26.00	0.644	0.887	0.282	0.129		24	8.26	4	23.69	0.525	0.879	0.409	0.161			
<i>Ac</i> 162.1	25	4.39	9	24.88	0.417	0.772	0.467	0.201		25	5.79	9	24.62	0.517	0.827	0.382	0.189			
Mean	28	11.46	8	28.24	0.533	0.884	0.405	0.187	0.393	28	11.39	8	27.38	0.542	0.893	0.400	0.190	0.344		
SE	3	2.14	1	3.02	0.051	0.024	0.050	0.023	0.015	3	1.91	1	2.89	0.030	0.017	0.028	0.017	0.024		

$N$ , sample size;  $N_A$ , mean number of alleles per locus;  $N_E$ , effective number of alleles;  $N_{PA}$ , mean number of private alleles;  $R$ , allelic richness;  $H_O$ , observed heterozygosity;  $H_E$ , expected heterozygosity;  $F_{IS}$ , inbreeding coefficient;  $r$ , estimated frequency of null alleles;  $F_{IS}$  (IIM), inbreeding coefficient accounting for null alleles; SE, standard error

### 3.2 Genetic clusters and Wahlund effects

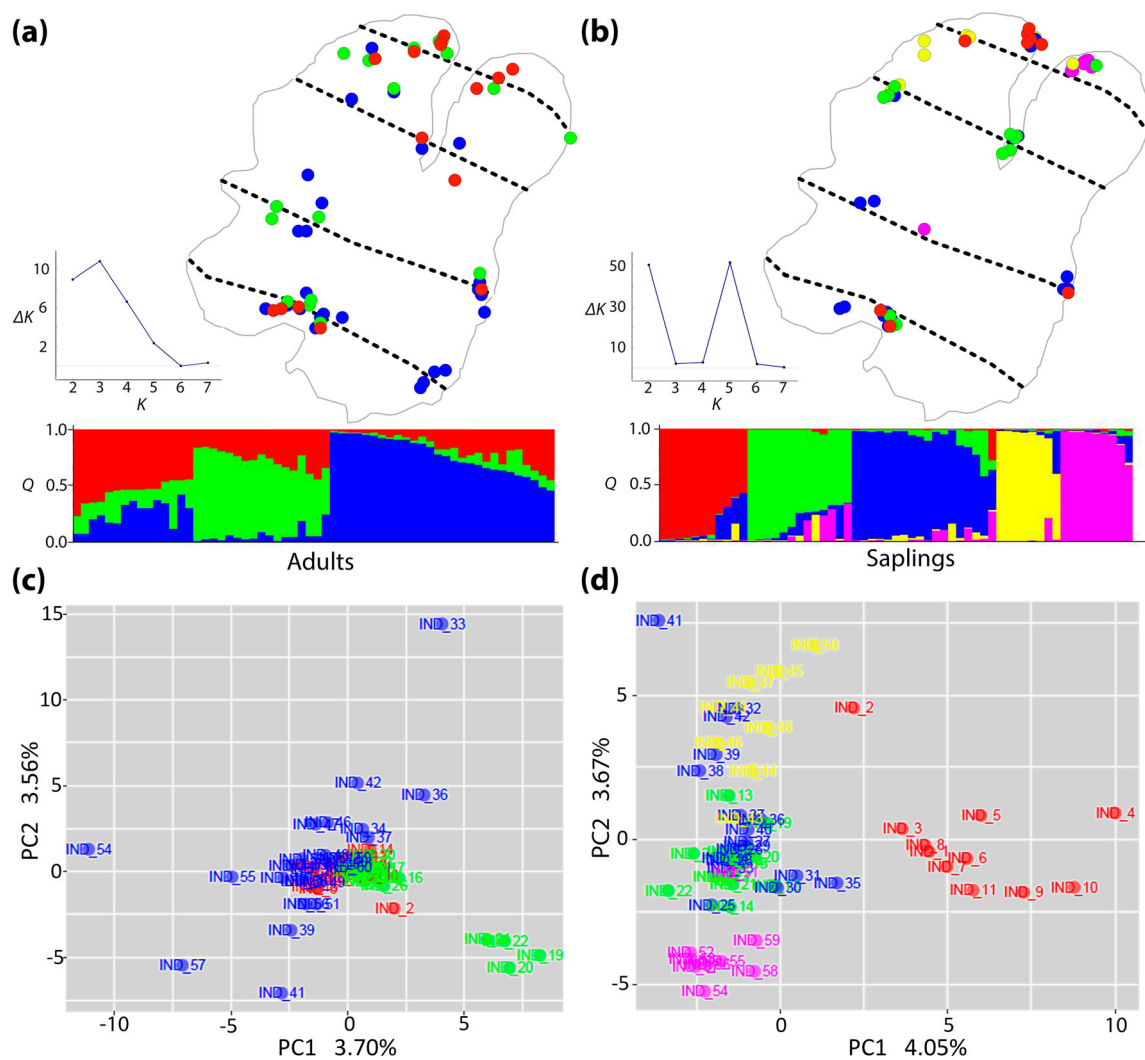
Bayesian model-based cluster analysis suggested  $K = 3$  and  $K = 5$  as the optimal numbers of clusters for adults and saplings, respectively (Fig. 2a and b). PCAs based on genotypes of adults and saplings separately also supported the existence of differentiated genetic clusters in the stand, in particular for saplings (Fig. 2c and d). Genetic clusters did not show any apparent spatial pattern (Fig. 2). As expected, the AMOVA indicated larger genetic variation within clusters than among them (Table 2). Nevertheless, genetic differentiation among clusters was significant ( $p < 0.05$ ). It was weak in adults ( $F_{ST} = 0.06$ ) and relatively high in saplings ( $F_{ST} = 0.14$ ), considering the short geographic distance between individuals from different genetic clusters (see Fig. 2a and b). Similar values were obtained when correcting for bias due to null alleles using the ENA method in FReeNA ( $F_{ST} = 0.05$ , CIs of 0.03–0.07, for adults, and  $F_{ST} = 0.12$ , CIs of 0.09–0.15, for saplings), indicating that null alleles had only a small effect on  $F_{ST}$  estimates. The harmonic mean of Jost's  $D$  (Jost 2008) allowed the evaluation of the relative differentiation of allele frequencies within and among clusters. This parameter was significant for both life stages and higher in saplings (0.70) than in adults (0.31, Table 2). This means that the three adult clusters shared more than half of the alleles, while the five sapling clusters had marked allelic differentiation and a low number of shared alleles. Finally, levels of inbreeding (as estimated by  $F_{IS}$ ) were still substantially high within each of the genetic clusters for both adults and saplings, allowing us to discard Wahlund effects (see Table 5 in Annex 2),

leaving assortative mating and selfing as the main factors to explain the high levels of inbreeding.

### 3.3 Fine-scale spatial genetic structure

The sPCA analysis revealed a significant global spatial genetic structure through the  $G$  test ( $p < 0.001$ ), but absence of local genetic structure (non-significant  $L$  test). The individual scores on the first eigenaxis of the sPCA indicated strong spatial autocorrelation, suggesting allele frequency gradients in both life stages that corresponded approximately with a geographic North-South axis of variation (Fig. 3). Allele frequency gradients were stronger in saplings (eig.sPCA = 0.178) than in adults (eig.sPCA = 0.136). Supporting these results, spatial autocorrelation analyses using SPAGeDi showed significant FSGS for both life stages ( $p < 0.001$ ), but with a more pronounced negative slope of pairwise kinship with distance for saplings, which was reflected in a 2.5-fold higher  $Sp$  statistic in saplings than in adults ( $Sp$  of 0.023 and 0.009, respectively; Table 3). Moreover, we found 5-fold greater pairwise kinship in the first distance class (up to 10 m) in saplings ( $F_1 = 0.328$ ) than in adults ( $F_1 = 0.059$ ), and the  $F_1$  of saplings was close to the value expected for full-sibs (i.e., 0.250) (Table 3, Fig. 4).

The MCMCglmm analysis (Table 6 in Annex 3) revealed that the model with a geographic predictor matrix alone had the lowest DIC and best explained the pattern of genetic relatedness in the stand; thus, altitude or DBH did not significantly contribute to explaining genetic relatedness.



**Fig. 2** Geographic representation of Bayesian model-based cluster analysis of population genetic structure (STRUCTURE software) for **a** adults ( $K = 3$ ) and **b** saplings ( $K = 5$ ), and principal component analyses (PCAs) based on **c** adult and **d** sapling genotypes.  $\Delta K$  vs.  $K$  distribution and the

bar plot ordered by the admixture proportions for each individual ( $Q$ ) are also shown in **(a)** and **(b)**. In **(c)** and **(d)**, STRUCTURE cluster membership is indicated using the same colors as in the bar plots **(a)** and **(b)**, respectively

### 3.4 Demographic history

Recent bottleneck signals were not detected in adults using Bottleneck program under a TPM model. On the contrary, statistically significant heterozygosity deficit ( $p < 0.05$ ) suggested a

recent population expansion. The Garza–Williamson  $M$  ratio, which typically detects older population size declines, showed also no evidence of bottlenecks ( $M = 0.37 \pm 0.12$ ). Based on ABC, demographic scenarios considering recent population decline/expansion had little support when compared to a null scenario of

**Table 2** Analyses of molecular variance (AMOVA), fixation indices, and allelic differentiation, for the genetic clusters of *Anadenanthera colubrina* var. *cebil* identified by STRUCTURE software

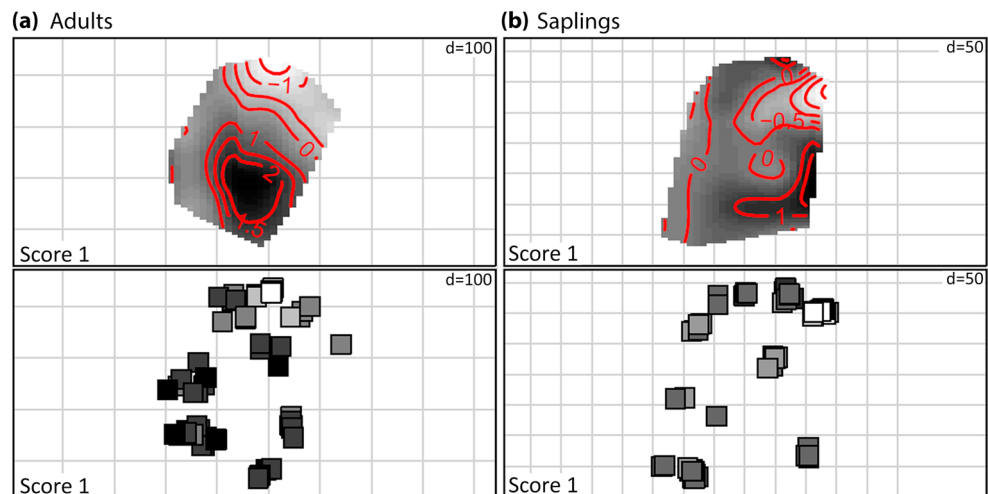
		Adults ( $K = 3$ )	Saplings ( $K = 5$ )
Percentage of variation	Among clusters	5.716	14.164
	Within clusters	94.284	85.836
Fixation indices	$F_{ST}^{(1)}$	0.0572*	0.1416*
	$F_{ST}^{(2)}$	0.0449 (0.0253–0.0673)	0.1188 (0.0897–0.1470)
Allelic differentiation	Jost's $D$	0.3092* (0.2454–0.3800)	0.6982* (0.6210–0.7707)

Values followed by \* were statistically significant ( $p < 0.05$ ); 95% confidence intervals (CIs) are indicated in parentheses, when available

<sup>(1)</sup> Standard estimates

<sup>(2)</sup> Null alleles were considered

**Fig. 3** Geographic representation of individual scores on the first sPCA axis of adults (a) and saplings (b), showing both raw values (lower panels) and their spatial interpolation (upper panels)



constant population size (Fig. 5, left). However, we found that an ancient expansion scenario was supported by high and consistent posterior probabilities (over 0.8; Fig. 5, right), with parameters: current effective population size ( $N_0$ ) of 7110 (CIs 3020–9900), ancestral population size ( $N_1$ ) of 712 (CIs 48–4360), and time of expansion ( $t$ ) of 357 (CIs 110–676) generations.

## 4 Discussion

Information on the historical and demographic events that shape the genetic structure of forest tree populations is relevant for understanding stand dynamics and their future evolution. We present here a case study that provides insights into the demographic population history of *A. colubrina* var. *cebil* in the Paranaense biogeographic province in Northern Argentina. Thirty years ago, the United Nations Food and Agriculture Organization (FAO 1986) reported that the species was suffering a slow decline. More recently, in the light of its ample use in the region, especially for extraction of wood and tannins, as well as for firewood (Justiniano and Fredericksen 1998; Tortorelli 2009), *A. colubrina* was indicated as a high priority species for in situ conservation (Monteiro et al. 2006). Although *A. colubrina* cannot be considered a threatened species (Roskov et al. 2017), sustainable forest management of its populations in the Paranaense biogeographic province demands attention, as most of them risk local

extinction, being located on private land that is exploited for logging.

In our study, we found high levels of genetic diversity and allelic richness in *A. colubrina* var. *cebil* from CSJ Nature Reserve, similar to those found in previous studies at wider geographical scales (Barrandeguy et al. 2014). High genetic diversity has also been detected in other Neotropical tree species (see Hardy et al. 2006) and is likely a consequence of life history traits of these trees, such as a long life span and a predominantly outcrossing mating system (Hamrick and Godt 1996; Petit and Hampe 2006). Despite the high levels of genetic diversity found in our study, we must be cautious when interpreting the intrinsic value of this remnant forest stand (Hedrick 1999). First, genetic diversity as estimated here only refers to neutral genetic diversity and not to variation associated with quantitative traits with adaptive value, and recent studies have shown little correlation between these two measures of diversity (e.g., Rodríguez-Quilón et al. 2015). Second, high genetic diversity is accompanied by high levels of inbreeding and differentiated genetic clusters, which also calls for caution in the use of this stand as seed source (Ellstrand and Elam 1993; Charlesworth 2003).

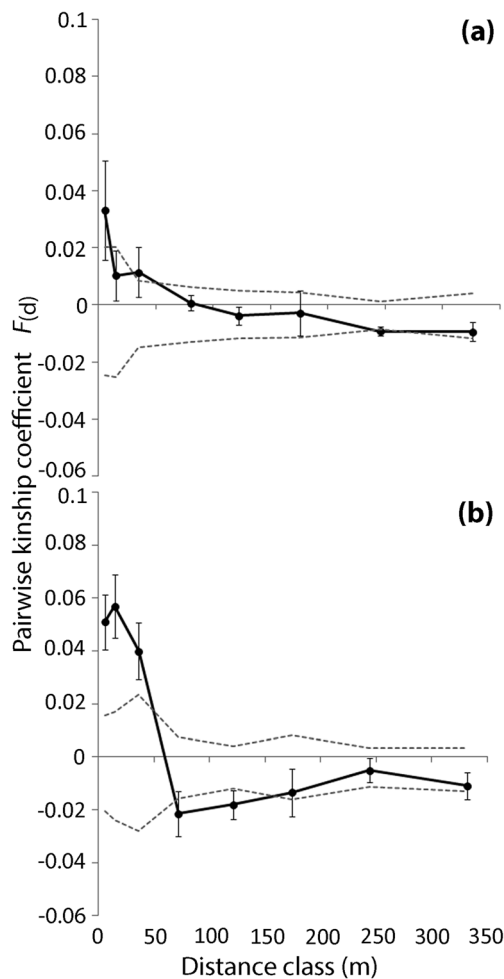
High inbreeding was detected for both adults and saplings at the stand level as well as within genetic clusters, even after accounting for null alleles [ $F_{IS}(IIM) = 0.40$  and  $0.34$ , respectively, at the stand level]. These inbreeding coefficients are substantially higher than those displayed by other insect-

**Table 3** Fine-scale spatial genetic structure (FSGS) in *Anadenanthera colubrina* var. *cebil*

Life stage	$N$	$F_1$	$b$	$Sp$
Adults	60	0.059*** (0.028)	− 0.0087*** (0.0029)	0.00921*** (0.00403)
Saplings	59	0.328*** (0.148)	− 0.0154*** (0.0028)	0.02295*** (0.01479)
Overall	119	0.205*** (0.082)	− 0.0074*** (0.0009)	0.00926*** (0.00171)

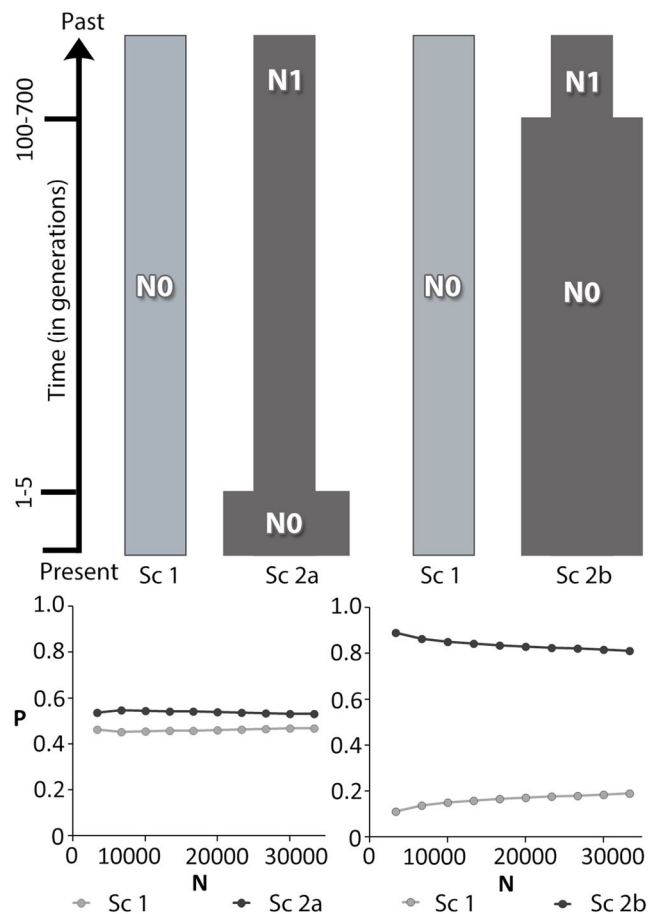
$N$ , number of individuals sampled;  $F_1$ , pairwise kinship in the first distance class (up to 10 m);  $b$ , regression slope of pairwise kinship with the logarithm of distance;  $Sp$ ,  $Sp$  statistic (see text). Values followed by \*\*\* were statistically different from zero ( $p < 0.001$ ); standard error (SE) is indicated in parentheses





**Fig. 4** Fine-scale spatial genetic structure (FSGS) correlograms for **a** adults and **b** saplings from the studied remnant population of *Anadenanthera colubrina* var. *cebil*; dashed lines indicate 95% confidence intervals per distance class around the hypothesis of random genetic structure obtained by permuting individual spatial locations as implemented in the SPAGeDi software

pollinated forest tree species such as *Dalbergia nigra* ( $F_{IS} = 0.08$ ), which is mainly pollinated by bees (Buzatti et al. 2012), *Cabralea canjerana* ( $F_{IS} = 0.06$ ) pollinated by moths (de Oliveira Melo et al. 2014), or *Prunus africana* ( $F_{IS} = 0.08$ – $0.19$ ), whose main pollinators are Hymenoptera and Diptera (Berens et al. 2014). Previous studies have also revealed a heterozygote deficit in *A. colubrina* var. *cebil* populations (Barrandeguy et al. 2014), although not as marked as in our study, and substantial levels of selfing and biparental inbreeding ( $t_m = 0.619$ ,  $t_m - t_s = 0.159$ ; Feres 2013). Together, these results suggest that *A. colubrina* var. *cebil* displays a mixed mating system. Selfing is common in legumes, and studies on the proportion of selfed vs. outcrossed offspring in other Neotropical legume trees have revealed mixed mating systems, for instance in *Senna multijuga* (Ribeiro and Lovato 2004), *Copaifera langsdorfii* (Tarazi et al. 2013), and *Dipteryx alata* (Tambarussi et al. 2017). More detailed investigation on the genetic constitution of offspring would be



**Fig. 5** Comparison of demographic scenarios (Sc) using ABC and considering: (1) constant population size, (2a) a recent population size change, and (2b) an ancient population size change. Posterior probabilities (P) of each scenario comparison (1 vs. 2a and 1 vs. 2b) are also shown

needed for a detailed characterization of the mating system of *A. colubrina* var. *cebil*, as well as its variation among populations. Nevertheless, an approximate selfing rate,  $s$ , can be estimated from the inbreeding coefficient using  $s = 2F_{IS}/(1 + F_{IS})$  (Hartl and Clark 2007), which suggests 51–56% selfing in the CSJ population. When species evolve to high proportions of selfing, this can decrease their effective population size, promote the accumulation of deleterious mutations, and hinder adaptation (Charlesworth 2003; Wright et al. 2013; Barrett et al. 2014). Small populations become inbred more rapidly than large populations, which, besides the mating system, is due to increased genetic drift and biparental inbreeding (Ellstrand and Elam 1993).

Knowledge on the influence of ecological context and life history traits on FSGS can provide crucial information for conservation of forest tree species. The mating system has been shown to have a strong effect on FSGS, with the highest  $Sp$  described in selfing species and, on average, a three times higher  $Sp$  ( $Sp = 0.037$ ) observed in mixed mating vs. outcrossing species ( $Sp = 0.013$ , Vekemans and Hardy 2004). The  $Sp$  statistic of *A. colubrina* var. *cebil* ( $Sp = 0.023$  and  $0.009$ , for saplings and

adults, respectively) was close to other Neotropical tree species with mixed mating systems, e.g., *Theobroma cacao* ( $Sp = 0.018$ , Silva et al. 2011) or *Dicorynia guianensis* ( $Sp = 0.026$ ; Hardy et al. 2006). The scale of pollen and seed movement also largely determines the level of kinship ( $F_{ij}$ ) observed among neighboring trees (Vekemans and Hardy 2004; Epperson 2003). In a recent meta-analysis for Neotropical tree species, Lowe et al. (2018) found that abiotically dispersed and pioneer species, such as *A. colubrina* var. *cebil* (Justiniano and Fredericksen 1998), have stronger FSGS than late successional species, whose seeds are dispersed by biotic vectors (see also Hardy et al. 2006).

In plant species, pollen and seeds often disperse at different scales (Anderson et al. 2010). In the case of *A. colubrina* var. *cebil*, while small bees are probably associated to pollen movement (e.g., *Tetragonisca angustula*; Flores and Sánchez 2010), seeds rely on autochory for dispersal. Flight distances for small bees can be as large as 621–951 m (Araújo et al. 2004), and thus they are probably a main agent effectively connecting *A. colubrina* var. *cebil* patches. Restricted seed dispersal can generate spatial clustering of full- and half-sibling cohorts localized close to the maternal plant (Nason and Hamrick 1997; Kalisz et al. 2001). This process, together with selfing, could then explain the high pairwise kinship in the first distance class for *A. colubrina* var. *cebil* saplings ( $F_1 = 0.328$ ), which was intermediate to the value expected for full-sibs (0.250) and selfs (0.500). High pairwise kinship values were not maintained in adults ( $F_1 = 0.059$ ). In addition, FSGS in adults was much weaker than in saplings ( $Sp$  of 0.009 vs. 0.023, respectively). While seed dispersal conditions the spatial configuration of sites available for establishment, early seedling recruitment represents a major demographical filter in the plant life cycle (Hampe et al. 2010). Thus, FSGS tends to decrease with age due to density-dependent mortality (Doligez et al. 1998; Epperson 2003; Leblois et al. 2004) and random demographic thinning (Hamrick et al. 1993; Schroeder et al. 2014), by which only few seedlings survive to become adults. The reproduction of *A. colubrina* var. *cebil* is characterized by the production of many seeds with low dormancy, high germination rates ( $\sim 70\%$ ) (Soldati and de Albuquerque 2010; de Souza et al. 2015), and high mortality rates of seedlings ( $\sim 50\%$ ) due to abiotic factors, competition, and soil microorganisms (de Medeiros et al. 2016). Therefore, the concerted effects of density-dependent mortality (e.g., Janzen–Connell effects; see Comita et al. 2014) and random demographic thinning could indeed have resulted in a more regular spatial distribution pattern and in a decrease of the FSGS from younger to older life stages in the studied stand.

In addition to significant FSGS, the existence of differentiated genetic clusters within the stand that are not explained by distribution in space alone is striking. Genetic differentiation among clusters was moderate, but still remarkable given the short separation distance between them ( $F_{ST}$  of 0.12 for

saplings and of 0.05 for adults, after correction for null alleles). A similar genetic structure was found in African populations of *Symphonia globulifera*, an ancient tropical tree species (Torroba-Balmori et al. 2017). Some genetic differentiation at local scales has also been reported for Neotropical forest tree species, e.g., in *Dicorynia guianensis* ( $F_{ST} = 0.02–0.03$ ), whose indehiscent flat pods are dispersed by anemochory (Latouche-Hallé et al. 2003). While in *Dicorynia guianensis* these patterns are explained by spatial segregation, evidence for aggregation following habitat features was found in *Symphonia globulifera* (Torroba-Balmori et al. 2017). In our case, however, the stand is relatively homogeneous and we did not find any evidence of habitat features underlying differentiated genetic clusters. This fact, together with only weak spatial segregation of genetic clusters, suggests assortative mating as the origin of differentiated genetic clusters in *A. colubrina* var. *cebil* (e.g., due to phenological differences among trees in the stand). Nevertheless, genetic diversity and allelic richness within clusters were still high for both adults and saplings, and not significantly different across generations, suggesting enough reproductive trees contributing to each genetic cluster.

Anthropogenic impact is usually considered a relevant factor involved in SDTF fragmentation and could have affected the demographic history of forest tree populations in this range. Recent data also indicated that *A. colubrina* is ecologically more labile than previously thought (Särkinen et al. 2011). However, assuming a seed-to-seed generation time of 50 years in *A. colubrina* var. *cebil*, our ABC approach suggested an ancient expansion timed approximately 18,000 years BP as best demographic scenario. Bottleneck tests also failed to show any signature of reduction in effective population sizes in our case study population from the CSJ Nature Reserve. Thus, anthropogenic disturbance and recent fragmentation do not seem to have caused detectable genetic effects in the studied forest remnant. However, it is still to be proven that this particular case study is representative of other populations in the region, and to what degree human-induced forest fragmentation may have affected these patterns.

Research on SDTFs has blossomed in the last decade (Hughes et al. 2013), revealing the biological value of these forests in terms of their highly endemic flora (DRYFLOR 2016), as well as providing deep insights into the biogeography of this biome and the processes that have shaped the historical assembly of SDTF species diversity (Hughes et al. 2013). Despite the fact that *A. colubrina* var. *cebil* is a key SDTF indicator species and notwithstanding the known human-induced fragmentation of the SDTFs, the studied remnant population located in the CSJ Nature Reserve conserves high genetic diversity. According to palaeodistribution modeling and biogeographic studies (Werneck et al. 2011; Pennington and Laving 2016), SDTFs generally represent temporally stable habitats over large time scales that are

believed to have maintained stable population sizes through climatic fluctuations. Moreover, in long-lived forest trees, especially those with long-distance pollen dispersal, the overlapping generations and longevity may act as a “buffer” against the effects of short-term fragmentation (Jones and Hubbell 2006).

## 5 Conclusion

Knowledge on patterns of genetic structure within populations and demographic history can help to understand future population trajectories and establish genetic conservation strategies. Our study in a remnant forest of *A. colubrina* var. *cebil* in the south of the Paranaense biogeographic province (CSJ Nature Reserve) found a mixed mating system and high levels of genetic diversity. Despite the fragmentation of the SDTF in this region, we observed moderate genetic structure (involving both FSGS and differentiated genetic clusters) as well as signatures of historical demographic expansions since the late Pleistocene. Our results suggest long-term population viability for the studied forest remnant and the value of CSJ Nature Reserve as a genetic reservoir. However, the high inbreeding coefficients, probably due to selfing associated with an increase in the levels of mating between relatives due to assortative mating, while not detrimental to current variability, suggest caution when using this stand as a source of material for reforestation.

**Acknowledgements** We thank Entidad Binacional Yacyretá for granting access to Campo San Juan Nature Reserve. Fieldwork benefited from the assistance of M.R. Kostlin. A.L. Goncalves wishes to thank to Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) for providing her doctoral fellowship. This study is part of the PhD thesis of A.L. Goncalves at the Universidad Nacional de La Plata.

**Funding** Financial support was provided by PICTO UNaM 2011 N°0133 from “Agencia Nacional de Promoción Científica y Tecnológica” and “Universidad Nacional de Misiones” to M.V. García, and by PIP N° 112-2015001-00860CO from CONICET, to M.V. García. This work has also been supported by ANR’s “Investissement d’Avenir” grants (CEBA: ANR-10-LABX-25-01).

**Data availability** Genetic and population data are available from the FIGSHARE repository (Goncalves et al. 2018). Dataset not peer-reviewed. Goncalves et al. (2018). Data from “Demographic history and spatial genetic structure in a remnant population of the subtropical tree *Anadenanthera colubrina* var. *cebil* (Griseb.) Altschul (Fabaceae)”. V1. figshare. [Dataset]. <https://doi.org/10.6084/m9.figshare.7488650.v1>.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## Appendix

### Annex 1: Population size change scenarios compared by means of approximate Bayesian computation (ABC)

**Table 4** Prior distributions and parameters used for population demographic scenario modelling in DIYABC software

Scenarios	Parameter	
Recent population size change	N0	Uniform (10–10,000)
	<i>t</i>	Uniform (1–5)
	N1	Uniform (10–10,000)
	Mean $\mu$	Log uniform ( $1 \times 10^{-3}$ – $1 \times 10^{-2}$ )
	Mean <i>P</i>	Uniform (0.1–0.5)
Ancient population size change	N0	Uniform (10–10,000)
	<i>t</i>	Uniform (100–700)
	N1	Uniform (10–10,000)
	Mean $\mu$	Log uniform ( $1 \times 10^{-3}$ – $1 \times 10^{-2}$ )
	Mean <i>P</i>	Uniform (0.1–0.5)

N0, current population size; N1, ancestral population size; *t*, time of the population size change; Mean  $\mu$ , mean mutation rate; Mean *P*, mean parameter of the geometric distribution

Summary statistics included in the computations, for each population: NAL (mean number of alleles across loci), HET (mean gene diversity across loci), VAR (mean allele size variance across loci), and MGW (mean Garza & Williamson’s *M* index across loci)

## Annex 2: Genetic diversity and inbreeding for the genetic clusters of *Anadenanthera colubrina* var. *cebil* identified by STRUCTURE

**Table 5** Genetic diversity within the three adult and five sapling genetic clusters found in the studied remnant population of *Anadenanthera colubrina* var. *cebil*

	Genetic clusters	$N$	$N_A$	$N_E$	$R$	$H_E$	$H_O$	$F_{IS}$	$F_{IS}$ (IIM)
Adults	1	15	10.75	6.69	4.94	0.807	0.539	0.34	0.270 [0.06–0.38]
	2	17	10.13	7.01	4.75	0.762	0.485	0.37	0.311 [0.13–0.46]
	3	28	22.13	19.70	6.63	0.942	0.558	0.41	0.378 [0.30–0.45]
Overall		60	28.38	12.90	6.02	0.892	0.533	0.41	0.393 [0.35–0.43]
Saplings	1	11	7.75	7.15	4.46	0.710	0.510	0.29	0.144 [0.03–0.28]
	2	13	8.25	5.53	4.50	0.749	0.519	0.32	0.267 [0.13–0.39]
	3	18	16.63	15.53	6.14	0.905	0.616	0.33	0.253 [0.13–0.36]
	4	8	6.63	5.96	4.70	0.791	0.391	0.52	0.295 [0.00–0.50]
	5	9	6.13	5.44	4.24	0.747	0.583	0.23	0.070 [0.00–0.21]
Overall		59	27.75	12.8	6.07	0.901	0.542	0.40	0.344 [0.28–0.42]

$N$ , sample size;  $N_A$ , mean number of alleles per locus;  $N_E$ , effective number of alleles;  $R$ , allelic richness;  $H_E$ , expected heterozygosity;  $H_O$ , observed heterozygosity;  $F_{IS}$ , inbreeding coefficient;  $F_{IS}$  (IIM), inbreeding coefficient accounting for null alleles; 95% confidence intervals (CIs) in brackets

## Annex 3: Association between genetic relatedness and environment/phenotypic traits

**Table 6** MCMCglmm analysis of genetic relatedness in adult individuals of the studied stand. DIC, deviance information criterion. See Materials and methods for details

Model	DIC	Delta DIC	DIC weight
NULL	−4133.795	10.241	0.003
GEO	−4144.036	0.000	0.464
ALT	−4131.802	12.234	0.001
DBH	−4131.792	12.244	0.001
GEO_ALT	−4142.096	1.940	0.176
GEO_DBH	−4142.122	1.914	0.178



## Annex 4: Distribution of *Anadenanthera colubrina* in Latin America



**Fig. 6** Distribution of *Anadenanthera colubrina* (white symbols) and *Anadenanthera colubrina* var. *cebil* (green symbols). The study site (Campo San Juan) is indicated. Occurrence data were downloaded from

the Global Biodiversity Information Facility and data points with inconsistent coordinates were excluded. [www.gbif.org](http://www.gbif.org) (2nd July 2018) GBIF Occurrence Download <https://doi.org/10.15468/dl.anvkwm>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Aguilar R, Quesada M, Ashworth L, Herrerias-Diego YV, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Mol Ecol* 17(24):5177–5188
- Aldrich PR, Hamrick JL, Chavarriaga P, Kochert G (1998) Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Mol Ecol* 7(8):933–944
- Anderson CD, Epperson BK, Fortin MJ, Holderegger R, James P, Rosenberg MS, Scribner KT, Spear S (2010) Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol Ecol* 19(17):3565–3575
- Araújo ED, Costa M, Chaud-Netto J, Fowler HG (2004) Body size and flight distance in stingless bees (Hymenoptera: Meliponini): inference of flight range and possible ecological implications. *Braz J Biol* 64(3B):563–568
- Barrandeguy ME, Prinz K, García MV, Finkeldey R (2012) Development of microsatellite markers for *Anadenanthera colubrina* var. *cebil* (Fabaceae), a native tree from South America. *Am J Bot* 99(9): e372–e374
- Barrandeguy ME, García MV, Prinz K, Pomar RR, Finkeldey R (2014) Genetic structure of disjunct Argentinean populations of the subtropical tree *Anadenanthera colubrina* var. *cebil* (Fabaceae). *Plant Syst Evol* 300(7):1693–1705

- Barrett SCH, Arunkumar R, Wright SI (2014) The demography and population genomics of evolutionary transitions to self-fertilization in plants. *Philos Trans R Soc Lond Ser B Biol Sci* 369(1648): 20130344
- Bauni V, Homberg M (2015) Reserva Natural Campo San Juan. Fundación de Historia Natural Félix de Azara, Buenos Aires 108 p
- Bauni V, Capmourteres V, Homberg MA, Zuleta GA (2013) Distribution and status of the extant xenarthrans (Mammalia: Xenarthra) in the southern cone mesopotamian savanna, Argentina. *Edentata* 14:35–50
- Berens DG, Braun C, González-Martínez SC, Griebeler EM, Nathan R, Böhning-Gaese K (2014) Fine-scale spatial genetic dynamics over the life cycle of the tropical tree *Prunus africana*. *Heredity* 113(5): 401–407
- Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol Ecol* 5:453–455
- Buzatti RS, Ribeiro RA, Lemos Filho JPD, Lovato MB (2012) Fine-scale spatial genetic structure of *Dalbergia nigra* (Fabaceae), a threatened and endemic tree of the Brazilian Atlantic Forest. *Genet Mol Biol* 35(4):838–846
- Cabrera AL (1971) Fitogeografía de la República Argentina. *Bol Soc Argent Bot* 14(1–2):1–42
- Carrió E, Rosselló JA (2013) Salt drying: a low-cost, simple and efficient method for storing plants in the field and preserving biological repositories for DNA diversity research. *Mol Ecol Resour* 14(2):344–351
- Chakraborty R, Zhong Y, Jin L, Budowle B (1994) Nondetectability of restriction fragments and independence of DNA fragment sizes within and between loci in RFLP typing of DNA. *Am J Hum Genet* 55(2):391–401
- Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* 24(3):621–631
- Charlesworth D (2003) Effects of inbreeding on the genetic diversity of populations. *Philos Trans R Soc Lond Ser B Biol Sci* 358(1434): 1051–1070
- Chybicki JJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* 100(1):106–113
- Cialdella AM (2000) Flora Fanerogámica Argentina, Fascículo 67: Fabaceae Subfamilia Mimosoideae. Proflora-CONICET, Córdoba 10 p
- Clark DA, Clark DB (1984) Spacing dynamics of a tropical rain forest tree: evaluation of the Janzen-Connell model. *Am Nat* 124:769–788
- Comita LS, Queenborough SA, Murphy SJ, Eck JL, Xu K, Krishnadas M, Beckman N, Zhu Y (2014) Testing predictions of the Janzen-Connell hypothesis: a meta-analysis of experimental evidence for distance-and density-dependent seed and seedling survival. *J Ecol* 102(4):845–856
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 114:2001–2014
- Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ, Guillemaud T, Estoup A (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* 24(23):2713–2719
- Csilléry K, Blum MG, Gaggiotti OE, François O (2010) Approximate Bayesian computation (ABC) in practice. *Trends Ecol Evol* 25(7): 410–418
- Davies SJ, Cavers S, Finegan B, White A, Breed MF, Lowe AJ (2013) Pollen flow in fragmented landscapes maintains genetic diversity following stand-replacing disturbance in a neotropical pioneer tree, *Vochysia ferruginea* Mart. *Heredity* 115(2):125–129
- de Medeiros RLS, de Souza VC, Neto MAB, de Araújo L, da Silva Barbosa A, de Medeiros RLS (2016) Estrutura da regeneração natural de *Anadenanthera colubrina* em fragmento de brejo de altitude em Bananeiras, PB. *Pesq Flor Bras* 36(86):95–101
- de Noir FA, Bravo S, Abdala R (2002) Dispersal mechanisms in some woody native species of Chaco Occidental and Serrano. *Quebracho* 9:140–150
- de Oliveira Melo AT, Coelho ASG, Pereira MF, Blanco AJV, Franceschinelli EV (2014) High genetic diversity and strong spatial genetic structure in *Cabralea canjerana* (Vell) Mart (Meliaceae): implications to Brazilian Atlantic Forest tree conservation. *Nat Conserv* 12(2):129–133
- de Souza TV, Torres IC, Steiner N, Paulilo MTS (2015) Seed dormancy in tree species of the Tropical Brazilian Atlantic Forest and its relationships with seed traits and environmental conditions. *Braz J Bot* 38(2):243–264
- Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proc Natl Acad Sci* 91(8):3166–3170
- Dick CW, Hardy OJ, Jones FA, Petit RJ (2008) Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Trop Plant Biol* 1(1):20–23
- Doligez A, Baril C, Joly HI (1998) Fine-scale spatial genetic structure with nonuniform distribution of individuals. *Genetics* 148(2):905–919
- DRYFLOR (2016) Plant diversity patterns in neotropical dry forests and their conservation implications. *Science* 353(6306):1383–1387
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4(2): 359–361
- Ellstrand NC, Elam DR (1993) Population genetic consequences of small population size: implications for plant conservation. *Annu Rev Ecol Syst* 24(1):217–242
- Epperson BK (2003) Geographical genetics. Princeton University Press, Princeton 376 p
- Epps CW, Keyghobadi N (2015) Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Mol Ecol* 24:6021–6040
- Escudero A, Iriondo JM, Torres ME (2003) Spatial analysis of genetic diversity as a tool for plant conservation. *Biol Conserv* 113(3):351–365
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14(8):2611–2620
- Ewédjè EBK, Ahanchédé A, Hardy OJ (2017) Breeding system, gene dispersal and small-scale spatial genetic structure of a threatened food tree species, *Pentadesma butyracea* (Clusiaceae) in Benin. *Conserv Genet* 18:799–811. <https://doi.org/10.1007/s10592-017-0928-8>
- Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10(3):564–567
- Fageria MS, Rajora OP (2013) Effects of harvesting of increasing intensities on genetic diversity and population structure of white spruce. *Evol Appl* 6(5):778–794
- Falguera C, Faggi A, Homberg M, Bogan S, Bauni V (2015) La vegetación de Campo San Juan. In: Bauni V, Homberg M (eds) Campo San Juan. Fundación Félix de Azara, Buenos Aires, pp 53–67
- FAO (1986) Databook on endangered tree and shrub species and provenances. FAO, Rome 524 p
- Feres JM (2013) Diversidade genética, fluxo gênico e sistema de cruzamento de *Anadenanthera colubrina* (Vell.) Brenan e *Anadenanthera peregrina* (L.) Speg: duas espécies que ocorrem em alta densidade no interior do Estado de São Paulo. PhD Dissertation, University of São Paulo, Brazil, 162 p
- Flores FF, Sánchez YAC (2010) First results of botanical characterization of honeys produced by *Tetragonisca angustula* (Apidae,



- Meliponinae) in Los Naranjos, Salta, Argentina. *Bol Soc Argent Bot* 45(1–2):81–91
- Frankham R (2003) Genetics and conservation biology. *C R Biol* 326:22–29
- García MV, Balatti PA, Arturi MJ (2007) Genetic variability in natural populations of *Paspalum dilatatum* Poir analyzed by means of morphological traits and molecular markers. *Genet Resour Crop Evol* 54:935–946
- Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Mol Ecol* 10(2):305–318
- Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010) Calculations of population differentiation based on  $G_{ST}$  and  $D$ : forget  $G_{ST}$  but not all of statistics. *Mol Ecol* 19:3845–3852
- Goncalves A, Garcia MV, Heuertz M, Gonzalez-Martinez S (2018) Data from “Demographic history and spatial genetic structure in a remnant population of the subtropical tree *Anadenanthera colubrina* var. *cebil* (Griseb.) Altschul (Fabaceae)”. V1. figshare. [Dataset]. <https://doi.org/10.6084/m9.figshare.7488650.v1>
- Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered* 86:485–486
- Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Softw* 33(2):1–22
- Hampe A, El Masri L, Petit RJ (2010) Origin of spatial genetic structure in an expanding oak population. *Mol Ecol* 19(3):459–471
- Hamrick JL (2004) Response of forest trees to global environmental changes. *For Ecol Manag* 197(1):323–335
- Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philos Trans R Soc Lond B Biol Sci* 351(1345):1291–1298
- Hamrick JL, Murawski DA, Nason JD (1993) The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. *Vegetatio* 107(1):281–297
- Hardy OJ, Vekemans X (2002) SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 2(4):618–620
- Hardy OJ, Maggia L, Bandou E, Breyne P, Caron H, Chevallier MH, Doligez A, Dutech C, Kremer A, Latouche-Hallé CE, Troispoux V (2006) Fine-scale genetic structure and gene dispersal inferences in 10 Neotropical tree species. *Mol Ecol* 15(2):559–571
- Hartl DL, Clark AG (2007) Principles of population genetics, 4th edn. Sinauer Associates, Sunderland 653 p
- Hedrick PW (1999) Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution* 53(2):313–318
- Hughes CE, Pennington RT, Antonelli A (2013) Neotropical plant evolution: assembling the big picture. *Bot J Linn Soc* 171(1):1–18
- Jombart T (2008) ADEGENET: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24(11):1403–1405
- Jones FA, Hubbell SP (2006) Demographic spatial genetic structure of the Neotropical tree, *Jacaranda copaia*. *Mol Ecol* 15(11):3205–3217
- Jost L (2008)  $G_{ST}$  and its relatives do not measure differentiation. *Mol Ecol* 17(18):4015–4026
- Jost L, Archer F, Flanagan S, Gaggiotti O, Hoban S, Latch E (2018) Differentiation measures for conservation genetics. *Evol Appl* 11(7):1139–1148
- Justiniano MJ, Fredericksen TS (1998) Ecología y silvicultura de especies menos conocidas: Curupá – *Anadenanthera colubrina*. Proyecto BOLFOR, Santa Cruz 58 p
- Kalisz S, Nason JD, Hanzawa FM, Tonsor SJ (2001) Spatial population genetic structure in *Trillium grandiflorum*: the roles of dispersal, mating, history, and selection. *Evolution* 55(8):1560–1568
- Klitgård BB, Lewis GP (2010) Neotropical Leguminosae (Mimosoideae). In: Milliken W, Klitgård B, Baracat A (Eds) Neotropikey—Interactive key and information resources for flowering plants of the Neotropics. [www.kew.org/science/tropamerica/neotropikey/families/Leguminosae\\_\(Mimosoideae\).htm](http://www.kew.org/science/tropamerica/neotropikey/families/Leguminosae_(Mimosoideae).htm). Accessed 21 Jan 2019
- Latouche-Hallé C, Ramboer A, Bandou E, Caron H, Kremer A (2003) Nuclear and chloroplast genetic structure indicate fine-scale spatial dynamics in a neotropical tree population. *Heredity* 91(2):181–190
- Leblois R, Rousset F, Estoup A (2004) Influence of spatial and temporal heterogeneities on the estimation of demographic parameters in a continuous population using individual microsatellite data. *Genetics* 166(2):1081–1092
- Loiselle BA, Sork VL, Nason JD, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *Am J Bot* 82(11):1420–1425
- Lowe AJ, Breed MF, Caron H, Colpaert N, Dick C, Finegan B, Gardner M, Gheysen G, Gribel R, Harris JBC, Kremer A, Lemes MR, Margis R, Navarro CM, Salgueiro F, Villalobos-Barrantes HM, Cavers S (2018) Standardized genetic diversity-life history correlates for improved genetic resource management of Neotropical trees. *Divers Distrib* 24(6):730–741
- Mogni VY, Oakley LJ, Prado DE (2015) The distribution of woody legumes in neotropical dry forests: the Pleistocene Arc Theory 20 years on. *Edinb J Bot* 72:35–60
- Monteiro JM, Cecília de Fátima CB, de Albuquerque UP, de Lucena RFP, Florentino ATN, de Oliveira RLC (2006) Use and traditional management of *Anadenanthera colubrina* (Vell) Brenan in the semi-arid region of northeastern Brazil. *J Ethnobiol Ethnomed* 2(6):6
- Morand M-E, Brachet S, Rossignol P, Dufour J, Frascaria-Lacoste N (2002) A generalized heterozygote deficiency assessed with microsatellites in French common ash populations. *Mol Ecol* 11(3):377–385
- Nason JD, Hamrick JL (1997) Reproductive and genetic consequences of forest fragmentation: two case studies of neotropical canopy trees. *J Hered* 88(4):264–276
- Peakall PE, Smouse R (2012) GenAlEx 6.5: genetic analysis in excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539
- Pennington TR, Laving M (2016) The contrasting nature of woody plant species in different neotropical forest biomes reflects differences in ecological stability. *New Phytol* 210(1):25–37
- Pennington TR, Prado DE, Pendry CA (2000) Neotropical seasonally dry forests and quaternary vegetation changes. *J Biogeogr* 27(2):261–273
- Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annu Rev Ecol Evol Syst* 37:187–214
- Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *J Hered* 90(4):502–503
- Prado DE, Gibbs PE (1993) Patterns of species distributions in the dry seasonal forests of South America. *Ann Mo Bot Gard* 80(4):902–927
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155(2):945–959
- Quantum GIS Development Team (2017) QGIS geographic information system. Open Source Geospatial Foundation Project. <https://qgis.org>. Accessed 21 Jan 2019
- Ribeiro RA, Lovato MB (2004) Mating system in a neotropical tree species, *Senna multijuga* (Fabaceae). *Genet Mol Biol* 27(3):418–424
- Ridley M (2003) Evolution, 3rd edn. Blackwell Publishing, Oxford 751 p
- Rodríguez-Quilón I, Jaramillo-Correa JP, Grivet D, Majada J, Alía R, Plomion C, Vendramin GG, González-Martínez SC (2015) Local effects drive heterozygosity–fitness correlations in an outcrossing long-lived tree. *Proc R Soc B* 282:20152230
- Roskov Y, Zarucchi J, Novoselova M, Bisby F (2017) ILDIS world database of legumes. In: Roskov Y, Abucay L, Orrell T, Nicolson D, Bailly N, Kirk PM, Bourgoin T, DeWalt RE, Decock W, de Wever A, van Nieukerken E, Zarucchi J, Penev L (eds) Species

- 2000 and ITIS catalogue of life. Naturalis, Leiden Digital resource at [www.catalogueoflife.org/col](http://www.catalogueoflife.org/col). Accessed 21 Jan 2019
- Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol Ecol Resour* 8(1): 103–106
- Sánchez-Azofeifa GA, Portillo-Quintero C (2011) Extent and drivers of change of Neotropical seasonally dry tropical forests. In: Dirzo R, Young HS, Mooney HA, Ceballos G (eds) *Seasonally dry tropical forests, ecology and conservation*. Island Press/Center for Resource Economics, Washington DC, pp 45–57
- Särkinen T, Iganci JR, Linares-Palomino R, Simon MF, Prado DE (2011) Forgotten forests-issues and prospects in biome mapping using seasonally dry tropical forests as a case study. *BMC Ecol* 11(1):27
- Schroeder JW, Tran HT, Dick CW (2014) Fine scale spatial genetic structure in *Pouteria reticulata* (Engl.) Eyma (Sapotaceae), a dioecious, vertebrate dispersed tropical rain forest tree species. *Glob Ecol Conserv* 1:43–49
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18(2):233–234
- Silva CRS, Albuquerque PSB, Ervedosa FR, Mota JWS, Figueira A, Sebbenn AM (2011) Understanding the genetic diversity, spatial genetic structure and mating system at the hierarchical levels of fruits and individuals of a continuous *Theobroma cacao* population from the Brazilian Amazon. *Heredity* 106(6):973–985
- Soldati GT, de Albuquerque UP (2010) Impact assessment of the harvest of a medicinal plant (*Anadenanthera colubrina* (Vell) Brenan) by a rural semi-arid community (Pernambuco), northeastern Brazil. *Int J Biodivers Sci Ecosyst Serv Manage* 6(3–4):106–118
- Spiegelhalter DJ, Best NG, Carlin BR, van der Linde A (2002) Bayesian measures of model complexity and fit. *Proc R Soc B* 64(4):583–639
- Steiner JJ, Poklemba CJ, Fjellstrom RG, Elliott LF (1995) A rapid one-tube genomic DNA extraction process for PCR and RAPD analyses. *Nucleic Acids Res* 23(13):2569–2570
- Tambarussi EV, Sebbenn AM, Alves-Pereira A, Vencovsky R, Cambuim J, da Silva AM, de Moraes MA, de Moraes MLT (2017) *Dipteryx alata* Vogel (Fabaceae), a neotropical tree with high levels of selfing: implications for conservation and breeding programs. *Ann For Res* 60(1):243–261
- Tarazi R, Sebbenn AM, Kageyama PY, Vencovsky R (2013) Edge effects enhance selfing and seed harvesting efforts in the insect-pollinated Neotropical tree *Copaifera langsdorffii* (Fabaceae). *Heredity* 110(6):578–585
- Torroba-Balmori P, Budde KB, Heer K, González-Martínez SC, Olsson S, Scotti-Saintagne C, Casalis M, Sonké B, Dick CW, Heuertz M (2017) Altitudinal gradients, biogeographic history and microhabitat adaptation affect fine-scale spatial genetic structure in African and Neotropical populations of an ancient tropical tree species. *PLoS One* 12(8):e0182515
- Tortorelli LA (2009) *Maderas y Bosques Argentinos*. Orientación Gráfica Editora, Buenos Aires 515 p
- Troupin D, Nathan R, Vendramin GG (2006) Analysis of spatial genetic structure in an expanding *Pinus halepensis* population reveals development of fine-scale genetic clustering over time. *Mol Ecol* 15(12):3617–3630
- van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4(3):535–538
- Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* 13(4):921–935
- von Reis Altschul S (1964) A taxonomic study of the genus *Anadenanthera*. *Contrib Gray Herb Harvard Univ* 193:3–65
- Wagner HH, Fortin MJ (2013) A conceptual framework for the spatial analysis of landscape genetic data. *Conserv Genet* 14(2):253–261
- Werneck FP, Costa GC, Colli GR, Prado DE, Sites JW (2011) Revisiting the historical distribution of seasonally dry tropical forests: new insights based on palaeodistribution modelling and palynological evidence. *Glob Ecol Biogeogr* 20(2):272–288
- Williamson-Natesan EG (2005) Comparison of methods for detecting bottlenecks from microsatellite loci. *Conserv Genet* 6(4):551–562
- Wright SI, Kalisz S, Slotte T (2013) Evolutionary consequences of self-fertilization. *Proc R Soc B* 280(1760):20130133
- Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends Ecol Evol* 11(10):413–418